

# 基于实时背景提取的人群密度检测

蒋绪团, 李庆敏

(厦门大学 计算机科学系, 福建 厦门 361005)

**摘要:**提出了一种实时的背景提取和更新算法,并把它运用在人群密度检测中。在去除干扰帧的基础上,通过帧间差分判定前景和背景像素,并使用累积数组对背景像素进行累积,然后结合 Surendra 算法提取和更新背景。在人群密度检测中运用已提取的背景图可以得到当前帧的前景图像,然后根据前景图像中人群的像素和边缘估计出人群的密度,最后利用坐标直方图的方法得到聚集度高的人群区域。实验表明,该背景提取方法快速有效,用于人群密度检测效果较好。

**关键词:**背景提取;累积数组;坐标直方图;人群密度检测

中图分类号:TP311 文献标识码:A 文章编号:1009-3044(2010)15-3996-03

## Crowd Density Detection Based on Real-time Background Extraction

JIANG Xu-tuan, LI Qing-min

(Department of Computer Science, Xiamen University, Xiamen 361005, China)

**Abstract:** A new method of real-time background extraction is proposed in this paper, and applies this method to crowd density estimation. On the Basis of removing interference frame, we can judge foreground and background pixels by inter-frame difference method, and use accumulative array cumulative background pixels, then extract and update background combining with Surendra algorithm. In crowd density estimation, it can get foreground image using background image, then estimation crowd density using foreground pixels and edges, finally it get crowd region which aggregation degree is high by Coordinate histogram. The experimental results indicate that the proposed background extraction method performs fast and effectively and obtains good results in crowd density estimation.

**Key words:** background extraction; accumulative array; coordinate histogram; crowd density detection

随着人们社会活动更加频繁,大量人群聚集在公共场所的可能性加大,如果不对这些聚集人群进行监控,就容易出现一些事故,对这些场所进行监控就显得十分必要。因此,人群密度检测有着广泛的应用前景和研究价值。

传统的人群密度监控是通个过人工来完成的,由于是凭借个人经验来发出危险警告,所以带有比较浓厚的主观色彩,这样容易出现发现太晚或者误报的情况。随着计算机和数字图像处理技术的发展和广泛应用,智能人群密度监控成为可能的同时也越来越受到人们的重视。目前智能人群密度估计的主要方法有基于像素的统计的方法<sup>[1]</sup>、基于纹理分析的方法<sup>[2]</sup>和基于个体特征的分析方法<sup>[3-4]</sup>。

在公共场所中,除了对人群的密度做出估计外,还应该特别注意聚集人群,因为危险常常发生在这些聚集人群中间。本文主要采用基于像素的统计的方法,并通过直方图累积的方法得到聚集人群的区域。该方法对背景的提取要求比较高,如果能提取出好的监控场景的背景,将大大提高检测的准确性。常用的背景提取的方法有:中值法<sup>[5]</sup>、均值法<sup>[6]</sup>、Surendra 方法<sup>[7]</sup>;基于预测的方法,如卡尔曼滤波器方法<sup>[8]</sup>、维纳滤波器方法<sup>[9]</sup>等;基于统计的方法,如高斯分布模型<sup>[10]</sup>、混合高斯分布模型<sup>[11]</sup>等。在研究了现有的一些方法后,本文提出了一种实时的背景提取算法,在去除干扰图像的基础上,使用累积数组对相邻帧差分后的图像进行累积,并结合 Surendra 算法提取和更新背景。该方法实时性高,提取效果较好。

## 1 实时背景提取和更新算法

### 1.1 算法步骤

背景的提取和更新是人群密度检测的重要组成部分,准确的监控场景背景提取可以大大提高人群密度检测的准确性。由于目前的一些背景提取和更新方法如常见的平均法、中值法等,需要消耗较多时间,而且提取的效果不理想,于是本文提出了一种实时的背景提取和更新算法,它提取和更新背景耗时少,效果较好。其算法如下:

- (1) 获取视频帧序列  $F_i$ 。
- (2) 若为  $i=1$ ,则把它当成背景  $B_1$ ,加入到累加数组,并转(1),否则转(3)
- (3) 判断  $F_i$  是否为干扰图像。由于信号干扰或摄像机抖动而获取的图像应该剔除,这样对后面的提取背景十分有利。定义

$$NF_i = \begin{cases} 0 & \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} (F_i(x, y) - B_{i-1}(x, y)) < T \\ 1 & \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} (F_i(x, y) - B_{i-1}(x, y)) > T \end{cases} \quad (1)$$

收稿日期:2010-03-10

基金项目:厦门市科技计划项目,项目编号:3502Z20073010

作者简介:蒋绪团(1984-),男,安徽阜阳市临泉县人,厦门大学计算机应用技术硕士研究生,主要研究方向为计算机视觉;李庆敏(1985-),厦门大学计算机应用技术硕士研究生。

其中  $N, M$  为帧图像的长和宽,  $B_{i-1}$  为背景图像,  $T$  为阈值, 实验中取  $T=0.8*M*N$ 。若  $NF_i=1$  则返回(1)。

(4) 和前一帧图像进行差分, 并对差分后的图像进行二值化, 实验中阈值取 30, 然后进行膨胀处理, 处理后的图像记为  $DF_i$ , 加入膨胀就能够把运动区域的边缘考虑进去, 使得到的背景效果更好。

(5) 把背景像素加入到累加数组  $C[N][M]$ :

$$\begin{cases} C[x][y] = C[x][y] + F_i(x, y) \\ CT[x][y] = CT[x][y] + 1 \end{cases} \quad \text{if } DF_i(x, y) = 0 \quad (2)$$

其中  $CT[x][y]$  为计数数组, 表示对坐标  $(x, y)$  的像素值进行累加, 这样方便后面的求平均操作  $DF_i(x, y)=0$  表示  $F_i$  中的背景像素。

(6) 如果  $i > N_1$  则表示要重新更新背景, 此时, 置  $i=2$  并且更新数组  $C$  和  $CT$ :

$$\begin{cases} C[x][y] = C[x][y] / CT[x][y] \\ CT[x][y] = 1 \end{cases} \quad (3)$$

此时数组  $C$  中的值即为对应背景像素值。这样就得到了背景  $B_i$ 。

(7) 如果  $i > N_2$  则通过计算  $C[x][y] = C[x][y] / CT[x][y]$ , 就可以得到背景  $B_i$ 。否则采用 Surendra 算法提取背景。实验中  $N_2$  取 100; Surendra 算法是 Surendra 等人提出的背景提取方法, 算法使用相邻两帧差值和阈值的比较来确定运动区域。更新背景时, 运动区域的背景保持不变, 而非运动区域当成背景, 以一定的权值来进行更新, 经过一段时间的迭代后, 就可以得到背景。算法的步骤如下:

(1) 把第一帧图像  $F_0$  作为原始背景  $B_0$ ;

(2) 令迭代参数  $i=1$ ;

(3) 求相邻帧的差分的二值图像  $TB_i$ ;

$$TB_i(x, y) = \begin{cases} 1 & \text{abs}(F_i(x, y) - F_{i-1}(x, y)) \geq T \\ 0 & \text{abs}(F_i(x, y) - F_{i-1}(x, y)) < T \end{cases} \quad (4)$$

表达式中:  $F_i, F_{i-1}$  分别表示当前帧图像和前一帧图像,  $\text{abs}(F_i(x, y) - F_{i-1}(x, y))$  为帧间差分,  $T$  为阈值。

(4) 由二值图像  $TB_i$  更新背景;

$$B_i(x, y) = \begin{cases} B_{i-1}(x, y) & \text{if } TB_i(x, y) = 1 \\ \alpha I_i + (1 - \alpha) B_{i-1}(x, y) & \text{if } TB_i(x, y) = 0 \end{cases} \quad (5)$$

表达式中:  $B_{i-1}(x, y), TB_i(x, y)$  分别表示背景图像和二值图像在点  $(x, y)$  处的值,  $\alpha$  为更新速度系数, 其值根据背景本身的慢变速度而选取, 实验中取 0.1;

(5)  $i=i+1$ , 返回到(3), 进行迭代。

Surendra 算法提取背景在短时间内可以达到不错的效果, 它运行速度快, 实时性高, 并且实现简单。但是当提取背景时间间隔较长时, 由于是简单的由更新速度系数控制更新, 容易丢失较早的背景信息。本文的方法用到它短时间内运行效果好的优点, 在前期运用它进行背景更新, 后期用累加数组更新背景。整个背景提取和更新的流程图如图 1 所示。

## 1.2 实验结果与分析

为了验证本文提出的算法, 在不同场景下进行实验, 如图 2 所示, 其中第一行、第二行和第三行分别为场景 1、场景 2 和场景 3, 视频帧大小  $320 \times 240$ , 并与其他算法进行比较。从实验结果可以看出本文提出算法较传统的均值法和中值法更具优越性。在运算速度方面, 平均法与中值法在三种场景下平均耗时分别为 1.25s、2.37s, 而本文的方法耗时 0.19s, 具有更快的运算速度。

## 2 人群密度检测

### 2.1 算法步骤

在公共场所中, 除了对人群的密度做出估计外, 还应该特别注意聚集人群, 因为危险常常发生在这些聚集人群中间。整个场景的人群密度小并不能说明没有聚集人群存在。所以, 应该对聚集人群单独作为一个估计模块。因此本文运用坐标直方图来提取聚集人群区域, 具体算法如下:

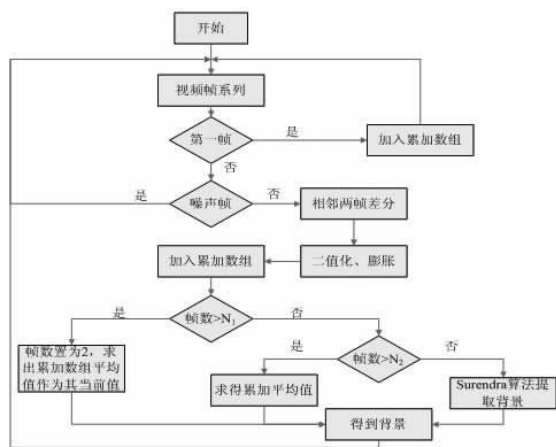


图1 背景提取和更新流程图

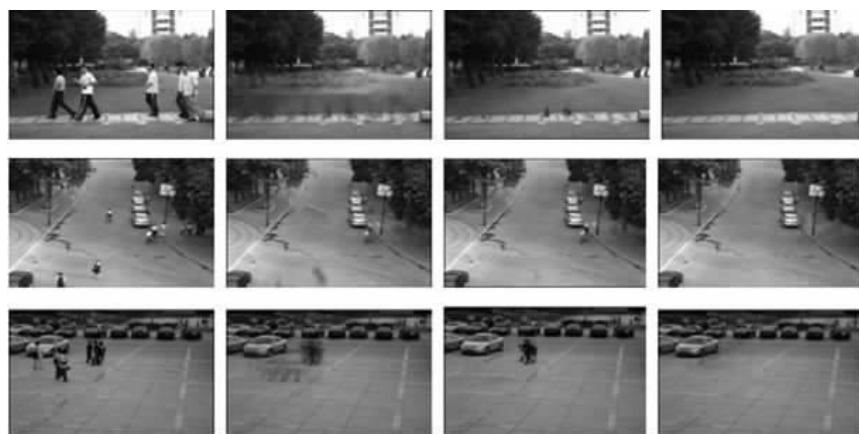


图2 在三种场景中不同算法提取的背景图像

- (1) 得到视频帧系列  $F_i$
- (2) 利用本文提出的背景提取算法,得到背景图像  $B_i$ ,并和当前帧进行差分,得到前景图像  $FW_i$ 。
- (3) 对前景图  $FW_i$  进行二值化,腐蚀,膨胀,得到图像  $TF_i$ 。
- (4) 对图像  $TF_i$  的横纵坐标分别进行直方图累积:

$$\begin{cases} CX[x]++ \\ CY[y]++ \end{cases} \quad \text{if } TF_i(x, y) = 1 \quad (6)$$

(5) 分别找出数组  $CX$  和  $CY$  的最大值  $x_{\max}, y_{\max}$ 。并以一定的步长  $\lambda$ ,如  $\lambda=5$  向四周生长,找到离  $(x_{\max}, y_{\max})$  最近的前景点  $(\tilde{x}, \tilde{y})$  作为下一步区域生长的种子点。

(6) 对  $TF_i$  进行膨胀 3 次,使聚集块相连。以  $(\tilde{x}, \tilde{y})$  为起始点,在  $TF_i$  上进行区域生长,得到聚集块的矩形区域。若矩形区域面积达到报警阈值  $T$ ,则标出矩形区域并报警。

## 2.2 实验结果与分析

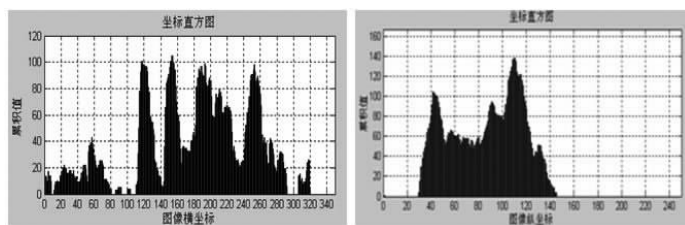
在本文提出的实时背景提取和更新算法的基础上根据上节的人群密度检测算法我们编程实现了人群密度检测系统,并经此子系统应用到视频监控系统中。图 3 是场景 1 中第 127 帧图像对应的坐标直方图,横坐标代表视频帧的宽度或高度,纵坐标代表前景区域像素坐标出现次数的累积值。如横坐标直方图里 (120, 100) 的点表示横坐标的值为 120 的前景像素总个数为 100。运用图 3(a)、3(b)横纵坐标直方图的最值  $M(154, 110)$  可以找到下一步区域生长的种子点  $S(156, 110)$ ,然后使用种子点  $S$  通过区域生长得到图 4(a)中红框标记的密集区域,4(b)是经过形态学处理后的前景二值图像。从图中可以看出,当场景中的人群密度达到一定的程度时系统就会框出警戒区域,并会发出警报。

## 3 结束语

本文提出的实时背景提取和更新算法与其他传统的算法相比具有快速有效的特点,在人群密度检测阶段本文主要采用基于像素的统计的方法,并通过坐标直方图的方法得到人群密度大的区域。通过实验可以看出将实时背景提取和更新算法应用到人群密度检测中,具有较好的检测效果。

## 参考文献:

- [1] Ma Ruihua, Li Liyuan. On pixel count based crowd density estimation for visual surveillance[C]. IEEE Conference on Cybernetics and Intelligent Systems, 2004: 170-173.
- [2] 郭森, 严和平, 柳伟. 基于灰度共生矩阵的密集人群人数估计[J]. 计算机工程与应用, 2008, 44(28): 169-171.
- [3] Valle J D, Oliveira L.E.S. People counting in low density video sequences[C]. LNCS 4872: PSIVT 2007, 2007: 737-748.
- [4] Li Chen, Ji Tao. People counting using iterative mean-shift fitting with symmetry measure [C]. International Conference on Information and Communication Security, 2007: 890-895.
- [5] Wu Bingfei, Juang Jhyhong, Tsai Pingsung, et al. A new vehicle detection approach in traffic jam conditions [J]. Computational Intelligence in Image and Signal Processing, 2007 (3): 1-6.
- [6] Fan Yong. A real-time algorithm of dynamic background extraction in image sequence [J]. Machine Learning and Cybernetics, 2005(8): 4997-5000.
- [7] Surendra Gupte, et al. Detection and Classification of Vehicles [J]. IEEE Transactions on Intelligent Transportation Systems, 2002, (3): 37-47.
- [8] Christof Ridder, Olaf MunKelf, Harald Kirchner. Adaptive background estimation and foreground detection using Kalman-Filtering [C]. International Conference on recent Advances in Mechatronics. Istanbul, Turkey: UNESCO Chair on Mechatronics, 1995. 193-199.
- [9] Kentaro Toyama, John Krumm, Barry Brumitt, et al. Wallflower: principles and practice of background maintenance [J]. Computer Vision, 1999, 1: 255-261.
- [10] Christopher Richard Wren, Ali Azarbayejani, Trevor Darrell, et al. Pfunder: real-time tracking of the human body [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 780-785.
- [11] Chris Stauffer, W.E.L. Grimson. Adaptive background mixture models for real-time tracking [J]. Computer Vision and Pattern Recognition, 1999, 2: 245-252.



(a) 横坐标直方图 (b) 纵坐标直方图

图3 坐标直方图



(a) 系统检测结果图 (b) 前景二值图

图4 人群密度检测结果图